

14: Hypothesis Tests

Stat 120 | Fall 2025

Prof Amanda Luby

```
library(tidyverse)
library(broom)
library(patchwork)
library(CarletonStats)
```

Note: This is the same **Pew** data that you looked at last week. Before starting, I recommend reviewing the question of interest and EDA that you created.

The data set **Pew.csv** contains part of a survey conducted by the Pew Research Center in January 2014. One of the questions they asked was, “Overall, when you add up all the advantages and disadvantages of the internet, would you say that the internet has been mostly a good thing, a bad thing, or some of both?”

The variable **values** codes the response as “good” if the respondent said the internet has been a good thing and “bad” otherwise (this includes “a bad thing” and “some of both”).

We want to formally test if this differs based on whether the respondent is 50 years or older or not.

```
Pew <- read.csv("http://math.carleton.edu/Stat120/RLabManual/Pew.csv")

Pew = Pew %>%
  mutate(
    over = ifelse(age >=50, "over50", "under50"),
    over2 = ifelse(age >=50, 1, 0)
  )
```

1. Write out appropriate null and alternative hypotheses.
2. Use the **permTest()** command in R to test your hypotheses. (Make sure to include **seed!**) You should be able to report the *statistic* and *p-value* from the R output. Then, make a *formal statistical decision* and *report your conclusion in context*.
3. Is this consistent with your results from last week?

4. We could have instead stated “We want to formally test if people under 50 have a more favorable view of the internet than people over 50”. How would the null and alternative hypothesis change?

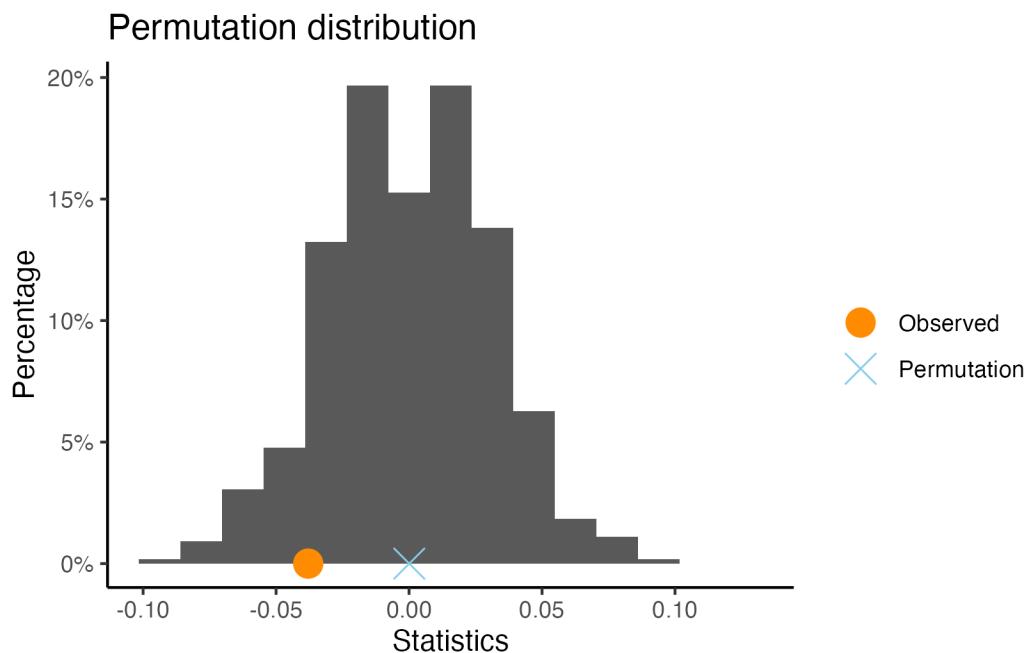
5. The code below runs a *one sided* test. Talk with your group to figure out if this is the correct test for your hypotheses in (4). If not, provide a correction. When you think you have the correct test, report the *statistic* and *p-value* from the R output. Then, make a *formal statistical decision* and *report your conclusion in context*.

```
permTest(values2 ~ over, data = Pew, seed = 10152024, alternative = "less")
```

** Permutation test **

```
Permutation test with alternative: less
Observed statistic
  over50 : 0.7808219    under50 : 0.8187311
Observed difference: -0.03791

Mean of permutation distribution: 7e-05
Standard error of permutation distribution: 0.03043
P-value: 0.1252
```



When you're done, please let me know!

If you have time, try replicating these results with StatKey. It won't be a built-in dataset, but you should be able to use the "edit data" button